Jet Algorithms at DØ

Elizabeth Gallas for the DØ Collaboration

1   Introduction

The DØ experiment at the Fermilab Tevatron proton-antiproton collider ($\sqrt{s} = 1.8$ TeV) accumulated a large sample of high energy jet production data during Run 1 (1992-1996). Since March 2001, DØ has engaged in continuous data collection with an upgraded detector equipped for the higher energy ($\sqrt{s} = 1.96$ TeV) and luminosity conditions of the Run 2 Tevatron. We summarize here pivotal measurements of Run 1 and consider their comparison to theoretical predictions and to other experiments. These factors elucidate DØ's jet clustering algorithm strategy for Run 2 jet measurements. Preliminary measurements of jets from the Run 2 DØ experiment are also presented.

2   The Measurement of Jets

Jet measurements described here rely primarily on DØ's finely segmented and calibrated calorimeter [1] to identify jets and measure their energy and orientation. The manner in which calorimeter cells are combined to define jets depends on the choice of jet clustering algorithm. DØ employed two such algorithms in Run 1: a cone algorithm and a 'k_T-type' algorithm.h This presentation focuses specifically on the Run 1 measurement of the inclusive jet cross section using both algorithms to explain DØ's jet clustering strategy for Run 2. The inclusive jet cross section measurement was chosen specifically because it is the most comprehensive and sensitive measurement of jets over the widest pseudorapidity (η) range and over the broadest range in jet transverse energy (E_T).

For the Run 1 measurements, jets were reconstructed using cone or k_T algorithms as indicated. Vertex, jet, and event quality criteria reduce backgrounds caused by electrons, photons, noise and cosmic rays. A jet "energy scale" corrects for calorimeter response, noise, showering outside of the cone radius (cone algorithm), and energy deposits from spectator interactions. Unsmearing corrections remove the effect of a finite jet energy resolution. Further details may be found in the publications cited in this paper.

3   Run I Cone Algorithm

The Run 1 cone algorithm reconstructs jets using an iterative algorithm with a fixed cone size of radius $R = 0.7$ in $\eta - \phi$ space. The pseudorapidity is defined as $\eta = -\ln[\tan\frac{\theta}{2}]$, where θ is the angle of the jet relative to the incoming proton beam; the angle ϕ is the azimuthal angle about the beam axis. E_T weighted centroids, initially centered on a cone axis corresponding to an energetic seed tower, are computed for the cells in the cone about the cone axis, iterating toward a cone axis which coincides with the E_T weighted centroid. Difficulties with a cone type algorithm include a seed ambiguity (a bias due to the choice of the most energetic towers as the initial jet centroids) and a split-merge ambiguity (the difficulty in defining when two nearby clusters of energy should be combined to form a single jet). The latter ambiguity necessitates the introduction of an R_{sep} parameter in the theoretical comparison: a minimal angular separation of two partons, mimicking the behavior observed at the calorimeter level, before those partons are considered to be distinct. The former ambiguity exacerbates the fact that the algorithm at higher orders in pQCD is not infrared safe (exhibits sensitivity to soft radiation) [2]. Nonetheless, most jet analyses used the cone algorithm to reconstruct jets in Run 1 because it was the first algorithm implemented (arising from the historic *Snowmass Accord* [3]), because it was more easily employed in the software level of the trigger, and because its systematic uncertainties became better understood earlier in the course of the analysis of the data.

Fig. 1 shows the Run 1 DØ measurement [4] of the η and E_T dependence of the inclusive jet production cross section using the cone algorithm and 95 pb^{-1} of integrated lumiosity. The solid curves show the theoretical
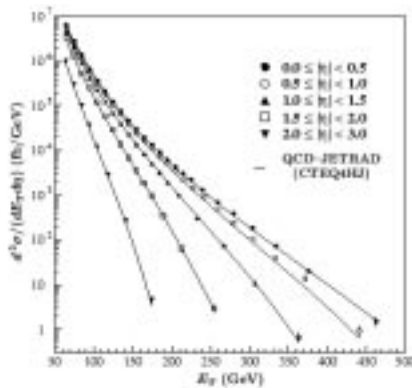

Fig. 1. DØ Run I inclusive cross section as a function of E_T in 5 η regions measured using the Run 1 cone algorithm.

prediction from JETRAD [5] using CTEQ4HJ PDF's. As indicated in the figure, and demonstrated in detail in a χ^2 analysis, the next-to-leading order predictions provide good overall agreement with the measurements. The jet energy regime and broad η reach of this measurement probes a significant area of previously unexplored phase space at high x ($10^{-3} < x < 1$) and Q^2 ($Q^2 > 2 \times 10^3 \text{GeV}^2$)

The precision of this measurement at high E_T has disappointingly ruled out new physics at this energy scale, but the results of this measurement refined parton distribution functions (PDFs) in this kinematic region. Specifically, the recent CTEQ6M and MRST2001 PDF sets incorporate these new results [6,7]. With these and many other jet measurements of Run 1, DØ brought the Tevatron into a new era of precision jet physics where for the first time, the uncertainty in the measurement is less than the theoretical errors.

4   Run I k_T Algorithm and Comparisons

In Run 1, DØ also explored the use of a recombinant or 'k_T-type' algorithm for jet reconstruction. The k_T algorithm used by DØ [8] combines energy clusters (i and j) based on their relative angular separation (ΔR_{ij} in $\eta - \phi$ space) and transverse energy ($E_{T,i}$ and $E_{T,j}$) by successive combination: $d_{ij} = min(E_{T,i}, E_{T,j})\Delta R_{ij}^2/D^2$ where D is a stopping parameter ($D = 1.0$ for this analysis) which approximately characterizes the size of the resulting jets. Clusters are combined, recalculating all $E_{T,i}$ and ΔR_{ij} until no d_{ij} is less than d_{ii}. The algorithm starts with a collection of pre-clusters (required to reduce the event size during data processing) separated by $\Delta R_{ij} \geq 0.2$. After running the successive combination algorithm, a list of jets is produced, each with transverse energy E_T and separated by $\Delta R > D$ from any other jet. By employing sensible choices in preclustering, the advantage of the use of this algorithm is that the same algorithm can be applied to the theory as in the measurement, greatly facilitating theoretical comparisons (no ad-hoc parameterizations, like R_{sep}, need to be introduced).

Fig. 2 shows DØ's Run 1 measurement [9] of the E_T dependence of the inclusive jet production cross section in the central region ($|\eta| < 0.5$) using the k_T algorithm and 87.3 pb^{-1} of integrated luminosity. The cross section exhibits reasonable agreement with next-to-leading order QCD predictions except at low E_T where a more significant divergence is observed.

To further explore these differences, the circular data points of Fig. 3 show the fractional difference between this k_T inclusive jet cross section measurement and the NLO QCD predictions as a function of E_T. The square data points show the corresponding distribution for the Run 1 cone algorithm. The cone algorithm result is in excellent agreement with the theory in both shape and overall normalization. In contrast, the k_T fractional difference portrays a mismatch with respect to both, but because the uncertainties
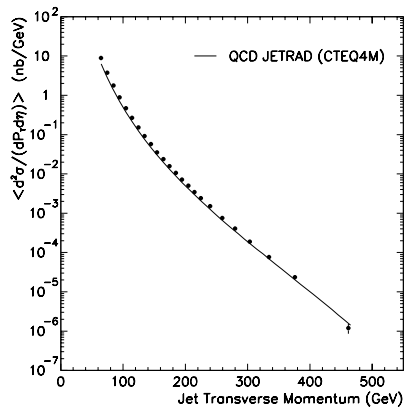
Fig. 2. DØ Run I inclusive cross section as a function of E_T in the central ($|\eta| < 0.5$) region measured using the k_T algorithm.
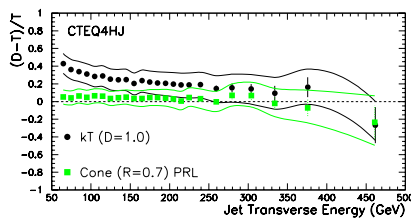

Fig. 3. Data - Theory / Theory for the inclusive jet cross section measured using k_T (circles) and cone (squares) algorithms. The JETRAD prediction assumes CTEQ4HJ PDF's.

are highly-correlated in E_T, the normalization differences are not remarkable. The departure in shape at low E_T in the k_T cross section, however, is not consistent with the uncertainties in the data. A full χ^2 analysis bears this out and moreover shows that if the first four data points are ignored, the k_T predictions are consistent with the observed cross section at the 77% probability level.

In a further analysis which matches jets found by the cone and k_T algorithm, it is found that the cross section for k_T differs from the cone result because the individual k_T jets in each event possess more energy than the matching cone jet. By removing the extra energy jet-by-jet, the energy difference accounts for the entire difference between the cross sections. Hadronization effects were explored in an attempt to account for the difference using

a particle level Monte Carlo simulation[10] and were found to account for possibly half of the E_T difference.

It is plausible that because jet phenomenology of QCD was largely developed within the context of a cone type algorithm, unexpected effects unique to the cone algorithm have been inadvertently incorporated in the PDFs. In that sense, agreement between theory and cone results is not surprising, nor is the marginal agreement between the predictions and the newer k_T algorithm results. Hadronization effects appear to represent at least some of this missing piece, but in their current form they are not enough to remove the observed discrepancy.

It is possible that a previously ignored missing effect is not modeled by the prediction and is exposed by the k_T algorithm. For this reason, it is important to continue to use the k_T algorithm in future analyses. Because of the marginal agreement with theory using the k_T algorithm, a cone type algorithm has been chosen by DØ as the primarly jet clustering algorithm but known limitations of the Run 1 cone algorithm must be addressed as discussed in the next section.

5   Run II

The collective experience of those involved with Run 1 jet measurements and theoretical predictions prove that consistent jet algorithm techniques between these groups facilitate comparison of those results. Many interested individuals are actively participating in a discussion/working group cast in 1999 to explore and define standard jet finding proceedures for all to use in Run 2. While the working group continues to discuss ongoing issues, many of their findings [11] have been adopted and incorporated into DØ's Run 2 analysis strategy.

In contrast to the E_T-weighted sums of calorimeter towers of the Run 1 cone algorithm, the Run 2 cone algorithm is defined by 4-momentum sums of the towers inside the cone. To eliminate sensitivity to infrared divergence of gluon radiation, effects of seed ambiguity are addressed by adding additional midpoint seeds between pairs of jets in close proximity. With these changes, DØ adopted the modified cone algorithm as the principle jet finding algorithm in Run 2, but DØ continues to explore the use of the k_T algorithm.

Figs. 4 and 5 show the preliminary Run 2 jet inclusive and dijet mass spectra obtained using the modified cone algorithm. Only statistical errors are shown and the corresponding data set is a fraction of the total collected to date. While the data have a preliminary energy correction $(30 - 50\%$ uncertainty), the plots lack important efficiency corrections and unsmearing of energy resolution effects so comparisons to theory would be premature.

From these plots, the extended kinematic reach of jet measurements in Run II is already apparent. With the increased center-of-mass energy and
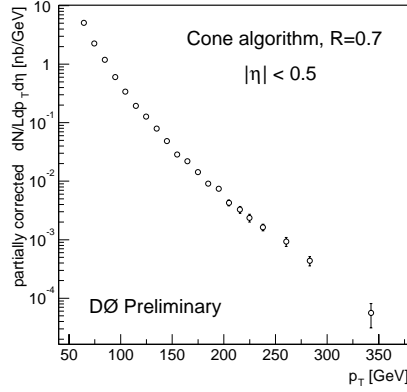
Fig. 4. Preliminary DØ Run 2 inclusive jet production cross section using the Run 2 cone algorithm.
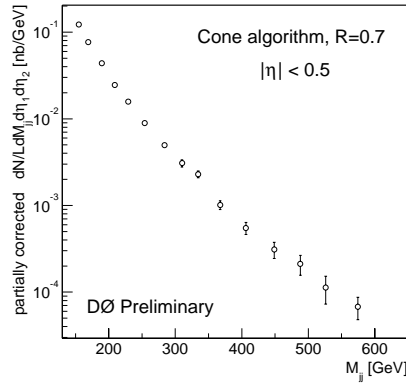

Fig. 5. Preliminary DØ Run 2 di-jet mass spectrum.

luminosity of the Tevatron and the capabilities of the upgraded DØ detector, we continue to explore the new energy frontier.

6   Acknowledgments

References

1. DØ collaboration, S.Abachi et al.: NIM A **338**, 185 (1994).
2. W.B. Kilgore and W.T. Giele: Phys. Rev. D **69**, 3615 (1992).
3. J.E. Huth et al.: in 'Proceedings of Research Directions For The Decade: Snowmass 1990', edited by E.L. Berger (World Scientific, Singapore, 1992) p134.
4. B. Abbott et al.: Phys. Rev. Lett. **86**, 1707 (2001).
5. W.T. Giele, E.W.N. Glover, David A. Kosower: Nucl. Phys. B **403**, 633 (1993).
6. J. Pumplin et al.: 'New Generation of Parton Distributions with Uncertainties from Global QCD Analysis', hep-ph/0201195
7. A.D. Martin et al.: Eur. Phys. J. C **23**, 73 (2002).
8. Frame K C: 'Implementation and Calibration of a k_T Jet Finding Algorithm for use in $p\bar{p}$ Collisions at $\sqrt{s} = 1.8$ TeV', Ph.D. thesis, Michigan State University, 1999.
9. V.M. Abazov et al.: Phys. Lett. B **525**, 211 (2002).
10. G. Marchesini et al.: Computer Phys. Commun. **67**, 465 (1992).
11. G.C. Blazey et al.: 'Run II Jet Physics', Proceedings of the Physics at RUN II: QCD and Weak Boson Physics Workshop, Batavia, Illinois, 4-6 Nov 1999, pp.47-77. (also FERMILAB-CONF-00-092-E, May 2000. 32pp or e-Print Archive: hep-ex/0005012)